Enhancing the Outcome of Microarray-Based Molecular Diagnostics

Tulpan, Dan; Léger, Serge

The Advances in Microarray Technology (AMT) [Proceedings], 2008

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

*Enhancing the Outcome of Microarry-Based Molecular Diagnostics **

Tulpan, D.C., Léger, S.
May 2008

* published at The Advances in Microarray Technology (AMT). May 7-8, 2008. Barcelona, Spain. NRC 50359. The Online Journal of Scientific Posters, 2008.

Canada

Enhancing the Outcome of Microarray-Based Molecular Diagnostics

Dan Tulpan, Serge Léger

Knowledge Discovery Group, Institute of Information Technology, National Research Council of Canada, Moncton, New Brunswick, Canada.

Abstract

Molecular diagnostics determines how genes and proteins are interacting in a cell and focuses upon gene and protein activity patterns in different types of cancerous or precancerous cells. Molecular diagnostics uncovers these sets of changes and captures this information as expression patterns via various data analysis techniques. The accuracy of the diagnostic process is hampered by many factors like: quality of the biological samples, accuracy of analytical methods and quality of microarray probes.

While the first two factors can be addressed by using better sample storage procedures and careful sample manipulation and by carefully choosing analytical techniques with wider domains of applicability and proven higher accuracies, the third factor is most of the time hard to adjust, given the rigidity of existing microarray platforms and increased costs for their re-design and re-validation.

Here, it is presented a technique that allows researchers to use experimental data (expression values) obtained with existing microarray platforms whose probe sets may not match the biological information known today. The technique is capable of adjusting the expression values in the experimental data by spotting inconsistencies between probes and their sequence origins and estimating their level of hybridization using molecular thermodynamic modeling.

Summary

Molecular diagnostics uncovers gene and protein activity patterns in different types of cancerous cells via data analysis techniques. The accuracy of the diagnostic process can be improved by adjusting the expression values in the experimental data using molecular thermodynamic modeling.

Speaker Biography

Dan Tulpan is a Research Officer at the Department of Information Technology of the National Research Council of Canada. He received a Ph.D. from the Computer Science Department of the University of British Columbia (Canada) in 2006 and was a Postdoctoral Fellow in Brinkman Laboratory at Simon Fraser University and instructor for the Masters of Software Systems Program at UBC. Dan's research interests are currently focused on topics of bioinformatics, biotechnology, artificial intelligence and empirical algorithms.

NRC·CNRC
Institute for Information Technology

National Research Council Canada | Conseil national de recherches Canada

Enhancing the Outcome of Microarray-Based Molecular Diagnostics

Dan Tulpan & Serge Léger
Knowledge Discovery Group, Institute of Information Technology, National Research Council of Canada, Moncton, New Brunswick, Canada

Background

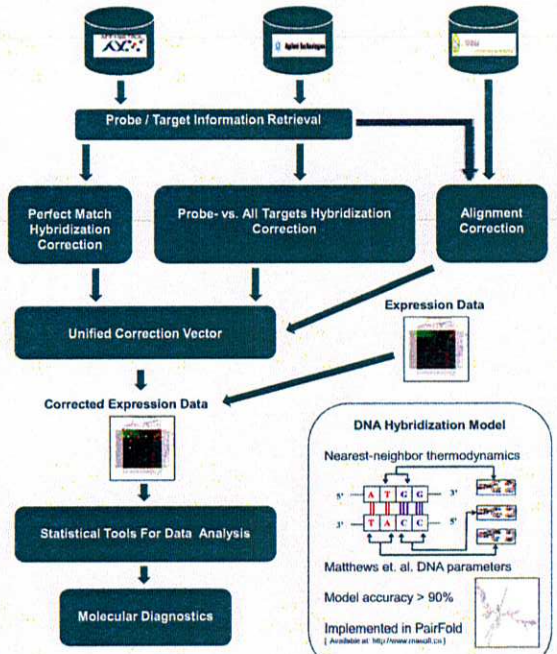
Molecular diagnostics determines how genes and proteins are interacting in a cell and focuses upon gene and protein activity patterns in different types of cancerous or precancerous cells. Molecular diagnostics uncovers these sets of changes and captures this information as expression patterns via various data analysis techniques. The accuracy of the diagnostic process is hampered by many factors like: quality of the biological samples, accuracy of analytical methods and quality of microarray probes.

While the first two factors can be dealt with by using better sample storage procedures and careful sample manipulation and by carefully choosing analytical techniques with wider domains of applicability and proven higher accuracies, the third factor is most of the time hard to adjust, given the rigidity of existing microarray platforms and increased costs for their redesign and revalidation.

Here, it is presented a technique that allows researchers to use experimental data (expression values) obtained with existing microarray platforms whose probe sets may not match the biological information known today. The technique is capable of adjusting the expression values in the experimental data by spotting inconsistencies between probes and their sequence origins and estimating their level of hybridization using molecular thermodynamic models.

Methods

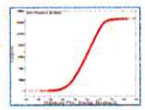
The following computational pipeline is proposed to adjust expression values based on platform dependent probe/target information and publicly available knowledge.



Results

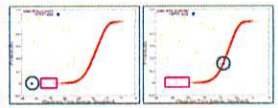
Perfect Match Hybridization Correction

Perfect match stabilities vary for different platforms, thus leading to different expression values and different sensing capabilities within the same platform.



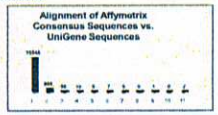
Probe vs. All Targets Hybridization Correction

Perfect match probes from different probe sets do not capture the correct target.

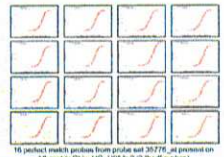


Alignment Correction

A simple sequence similarity check reveals a large number of one-to-many relationships between target sequences used in probe design and currently known sequence information.



For most perfect match probes there is a set of targets that can potentially hybridize with the probe, thus competing with the desired target (highlighted in purple).



Future Directions

A thorough evaluation, comparison and in vivo validation of the current approach is under investigation.

We are open for collaborations with academic and industrial partners. If you are interested in this technology please contact: Dr. Dan Tulpan at **dan.tulpan@nrc-cnrc.gc.ca**